# Object Tracking Based on PISC Image and Template Matching

Budi Sugandi
Electrical Engineering Department
Batam State Polytechnic
Batam, Indonesia
budi_sugandi@polibatam.ac.id

***Abstract*—This paper proposed a method for object tracking by using a peripheral increment sign correlation image and template matching. This method can be used to evaluate sign changes of the neighborhood brightness. The method is constructed from only the trend of the brightness changes in the neighborhood of the pixel under consideration. Consequently, the image is highly robust to brightness changes over the sequence, and the similarity of the original texture pattern can be detected even if there is a brightness change. In order to track the interest object among the objects detected in the field of view of camera, the template matching method was applied. At the last stage, we evaluated the proposed method by identifying the objects using their color and spatial feature. The experimental results showed that proposed method has identification rate more than 90%.**

***Keywords—object tracking; PISC image; template matching; color feature; spatial feature***

## I. Introduction

In recent years, with the latest technological advancements, visual surveillance and security system receive a great deal of interest. Until recently, video surveillance and security system was mainly a concern for military or large-scale companies. However, increasing crime rate, especially in metropolitan cities, necessitates taking better precautions in security-sensitive areas, like country borders, airports or government offices. Even individuals are seeking for personalized security systems to monitor their houses or other valuable assets. The sole answer for this increasing demand for personal and societal security is automation.

The broad range of applications motivates the interests of many researchers worldwide to develop and build the robust and reliable security system. Recent developments of security system are introduced by many researchers especially in the fields of detecting and tracking the moving objects.

Various methods for detecting and tracking of moving objects have been proposed in the past. Liu et al. [1] proposed a background subtraction to detect moving regions in an image by taking the difference between the current and the reference background image in a pixel-by-pixel fashion. It is extremely sensitive to change in dynamic scenes derived from lighting and extraneous events etc. Lipton et al. [2] proposed a frame difference technique that use of the pixel-wise differences between two or three successive frames in an image to extract moving regions. This method is adaptive to dynamic environments, but generally does a poor job of extracting all the relevant pixels, e.g., there may be holes left inside moving entities. Meyer et al. [3] proposed an optical flow method by computing the displacement vector field to initialize a contour based tracking algorithm, called active rays, for the extraction of articulated objects. The optical flow method can be used to detect independently moving objects even in the presence of camera motion. However, most flow computation methods are computationally complex and very sensitive to noise. In another work, Wren et al. [4] used statistical texture properties of the background observed over extended period of time to construct a model of the background, and use this model to decide which pixels in an input image do not fall into the background class. The fundamental assumption of the algorithm is that the background is static in all respects: geometry, reflectance, and illumination. Also Davis el al. [5] had approach that based upon image motion only presuming that the background is stationary or at most slowly varying, but that the person is moving. In these methods, no detailed model of the background is required. These methods are only appropriate for the direct interpretation of motion; if person stops moving, no signal remains to be processed. These methods also require constant or slowly varying geometry, reflectance and illumination.

To overcome those problems, we propose a new method for tracking of moving objects employing peripheral increment sign correlation (PISC) [6] image as an image matching method with robust performance, high accuracy, and high computational efficiency. The PISC image is constructed from only the trend of the brightness changes in the neighborhood of the pixel under consideration. The extraction method proposed in this paper focuses on discriminating between the similar background areas without being affected by brightness changes due to adverse conditions, by utilizing the robust matching performance of the increment sign correlation procedure. The proposed method is based on applying such a similarity decision to all pixels in the scene. In order to track the interest object among the objects detected in the field of view of camera, the template matching method was applied. At the last stage, we evaluated the proposed method by identifying the tracked objects using their color and spatial feature [7].

The rest of the paper is organized as follows. In section 2, we give a detail of our proposed method to detect and track the interest object. In section 3, the experimental and identification result of the proposed method are discussed. Finally, the conclusions and the future works are described in section 4.

## II. METHOD

We divided our proposed method into three stages; object detection, object tracking and object identification. On each stage, we apply our new method and evaluate the effectiveness of our method.

The first stage is done to detect the moving object emerging in the background by applying the PISC image. The second stage is performed to track the moving object using template matching. As the last stage, we identify the tracked object by extracting their color and spatial information of the object. In this paper, we use mean and standard deviation value as image feature of the tracked object. The details of each technique are described as follows.

### *A. PISC image*

The peripheral increment sign takes a value of 1 or 0 according to whether the increment near the considered pixel is positive or negative. This is a logical code representing the trend of brightness change. The increment sign is a code that represents the trend of the brightness change in a certain direction. The correlation coefficient over the whole image is called the increment sign correlation.

The PISC image is used to detect the moving person based on the trend of the brightness changes in the neighborhood of the pixel under consideration. The PISC image consider the brightness changes in the 16-neighborhood of the considered pixel as shown in Fig. 1. In order to detect the moving person in the image sequence, the background image is defined first.

In the background image $F = \{f_{i,j}\}_{j=1,\cdots,M-1}^{i=1,\cdots,N-1}$ the peripheral increment sign $b_{k(i,j)}$ is defined as,

$$b_{0(i,j)} = \begin{cases} 1 & (f_{i+2,j} \geq f_{i,j}) \\ 0 & (otherwise) \end{cases}$$
$$\vdots \tag{1}$$
$$b_{15(i,j)} = \begin{cases} 1 & (f_{i+2,j-1} \geq f_{i,j}) \\ 0 & (otherwise) \end{cases}$$

For any object image $G = \{g_{i,j}\}$ in the image time sequence, $b'_{k(i,j)} (k = 0,1,...,15)$ is similarly defined. In two images, the extent of matching between the increments signs in the 16 directions at the corresponding position is defined as the peripheral increment sign correlation $B$:

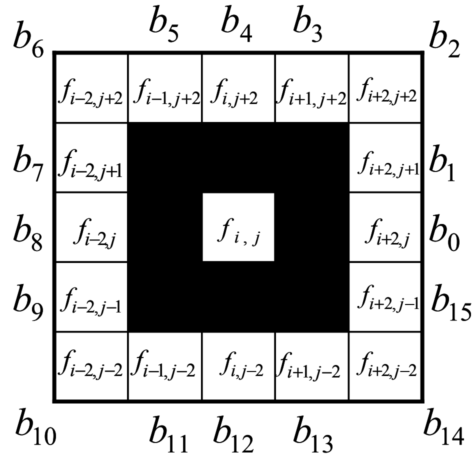



Fig. 1. Neighborhood for peripheral increment sign

$$B = \frac{1}{16}\sum_{k=0}^{15} c_{k(i,j)} \tag{2}$$

$$C_k = b_{k(i,j)} \bullet b'_{k(i,j)} + \left(1 - b_{k(i,j)}\right) \bullet \left(1 - b'_{k(i,j)}\right) \tag{3}$$

The value of B at each pixel is compared to some threshold and a decision is made whether it is similar or non-similar pixel. The PISC image is defined as binary image $\{I_{i,j}\}$ that defined as follows,

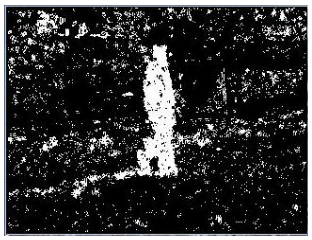
$$I_{i,j} = \begin{cases} 1 & (B \leq B_T) \\ 0 & (otherwise) \end{cases} \tag{4}$$

where $B_T$ is threshold usually takes a value in the range from 0.5 to 1. Fig. 2 shows an example of a PISC images in different threshold. Fig. 3 shows the comparison between background subtraction and PISC image. On that image, we can understand that background subtraction is very sensitive to change in dynamic scenes derived from lighting and extraneous events, while the PISC image can reduce the noise because of those effects.

In the next step, we performed a morphological operation [8] to reduce the image noises in the emerging object. The morphological operation implemented in this research is dilation followed by erosion. In dilation, each background pixel that is touching an object pixel is changed into an object pixel. Dilation will add pixels to the boundary of the object and close isolated background pixel. In erosion, each object pixel that is touching a background pixel is changed into a background pixel. Erosion will remove isolated foreground pixels. Morphological operation eliminates background image noises and fills small gaps inside an object. This property makes it well suited to our objective since we are interested in generating object masks which preserve the object boundary.

(a) Original image

(b) PISC image with threshold 0.5

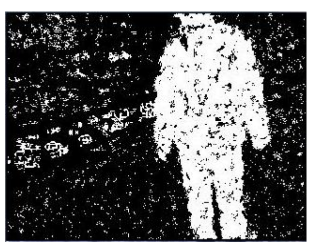


(c) PISC image with threshold 0.8

Fig. 2. Example of PISC image

(a) Original image

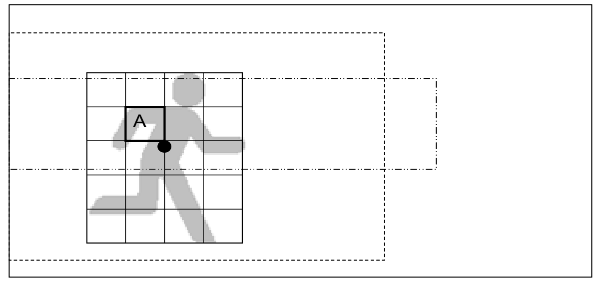
(b) Background subtraction



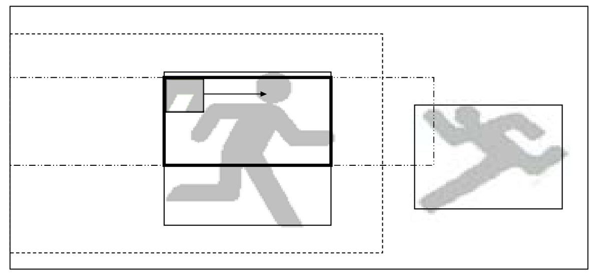
(c) PISC image

Fig. 3. Comparison between background subtraction and PISC image

Connected component labeling is performed to label each moving object emerging in the background. The connected component labeling groups the pixels into components based on pixel connectivity (same intensity or gray level) [9]. In this paper, connected component labeling is performed by comparing a pixel with the pixels in the four neighbors from top-left to bottom-right and from bottom-right to top-left.




(a) $(k\text{-}1)^{th}$ frame



(b) $k^{th}$ frame

Fig. 4. Template matching process

### B. Template Matching

In this paper, we use template matching technique to search the interest moving person among the objects appear on the scene. The algorithm of the template matching method is shown in Fig. 4. Firstly, to reduce the processing time, the template images are made only in the area of the extracted person of $(k\text{-}1)^{th}$ frame. The template size is 9×9. Then, each template image of $(k\text{-}1)^{th}$ frame search the target person in the area of the extracted person of $k^{th}$ frame. The correlation value between template image of $(k\text{-}1)^{th}$ frame and $k^{th}$ frame is evaluated by using the expression in (3). The matching template will have high correlation value while the non-matching template will have low correlation value. The numbers of the matching template of $(k\text{-}1)^{th}$ frame and $k^{th}$ frame are counted. The target person in $k^{th}$ frame is determined as a same person in $(k\text{-}1)^{th}$ frame when the numbers of matching templates are higher than the threshold value.

### C. Feature Extraction and Object Identification

Object identification is the last stage of our study. In this paper, the extracted features are divided into two types; color and spatial information of the moving objects.

The RGB color is used as color information of the moving objects. To obtain more color information for identification, we divide the moving object into three parts; the head, the upper and the lower part of the body. However, we only calculate the color information of upper and lower part of the human body.

The first color information calculated is mean value of each human body part as calculated by (5). The mean value is calculated for each color component of RGB space.

$$\mu_{f_k}^{O^i} = \frac{\sum_{x=x_{\min}^i}^{x_{\max}^i} \sum_{y=y_{\min}^i}^{y_{\max}^i} f_k(x,y)}{\#O^i} \quad (5)$$

where $i$ is number of the moving objects and $(x, y)$ is the coordinate of pixels in moving object. $\left(x_{max}^i, y_{max}^i\right)$ and $\left(x_{min}^i, y_{min}^i\right)$ are the maximum and minimum coordinates of moving object $i$, $f_k(x,y)$ denotes pixel value for each color component in RGB space of the current frame, $O^i$ denotes the set of the coordinate in the interest moving object $i$ and $\#O^i$ is the number of pixels of moving object $i$, respectively.

We can extract more useful color features by computing the standard deviation of each human body part as shown in (6).

$$SD_{f_k}^{O^i} = \sqrt{\frac{\sum_{x=x_{\min}^i}^{x_{\max}^i} \sum_{y=y_{\min}^i}^{y_{\max}^i} \left(f_k(x,y) - \mu_{f_k}^{O^i}\right)^2}{\#O^i}} \quad (6)$$

where, $\mu_{f_k}^{O_l^i}$ is the mean value and $SD_{f_k}^{O_l^i}$ is the standard deviation of each color component of the moving object, respectively.

The feature of objects extracted in the spatial domain is the position of the tracked object. The bounding box as defined in (7) is used as spatial information of moving objects.

$$\begin{aligned} B_{\min}^i &= \left\{\left(x_{\min}^i, y_{\min}^i\right) \mid x, y \in O^i\right\} \\ B_{\max}^i &= \left\{\left(x_{\max}^i, y_{\max}^i\right) \mid x, y \in O^i\right\} \end{aligned} \quad (7)$$

where, $B^i_{min}$ is the left-top corner coordinates and $B^i_{max}$ is the right-bottom corner coordinates, respectively.

After the extracted feature is obtained, we then calculate the similarity between the tracked object and the identified object as expressed in (8). The object with high similarity compared to certain threshold shows the similar object to the identified object, otherwise it will identify as different object.

$$\begin{aligned} S\left(F^i, F^j\right) &= Mc\left(\left|\mu_{f_k}^{O_l^i} - \mu_{f_k}^{O_l^j}\right|\right) + Mc\left(\left|SD_{f_k}^{O_l^i} - SD_{f_k}^{O_l^j}\right|\right) \\ &+ 0.5Mp\left(\left|B_{\min}^i - B_{\min}^j\right|\right) + 0.5Mp\left(\left|B_{\max}^i - B_{\max}^j\right|\right) \end{aligned} \quad (8)$$

where $Mc$ and $Mp$ are membership function for color and spatial information.

## III. EXPERIMENTAL RESULTS

We have done the experiments by using a camera in outdoor environments under noisy background and real time condition over $320 \times 240$ pixels image. We used the template size of $9 \times 9$ pixels. The template image was made from the object appear for the first time. We track the interest moving object from two and three moving objects appear in the background. The experimental results are shown in Fig. 5 - Fig. 7. The rectangle area on the object shows the extracted moving object. On each experiment, we apply the identification process using color and spatial feature of the moving object. The identification result is shown in table 1.

For the first experiment as shown in Fig. 5, two moving objects are moving in the different direction. At first, the man wearing the white shirt enters the scene from the left side. This person will be tracked as the interest moving object. Then on the next frame, the man wearing the blue shirt enters the scene from right side. They move in the different direction and overlap each other in the middle of the scene. We successfully track the first moving object as the interest moving object as our assumptions.

In the second experiment as shown in Fig. 6, two moving objects are moving in the same direction. At first, the man wearing the blue shirt enters the scene from the left side. This person will be tracked as interested moving person. Then on the next frame, the man wearing the white shirt enters the scene from left side also. They move in the same direction and overlap each other. We successfully track the first moving object as the interest moving object.

The last experiment as shown in Fig. 7, we track the interest moving object among three moving objects appear in the background. We also successfully track the first moving object as interested moving object.

## IV. CONCLUSION AND FUTURE WORKS

This paper proposed a method for detecting and tracking the moving object based on PISC image and template matching for real time application and an identification method using color and spatial feature. By using our proposed method, the satisfactory results are achieved. The tracking results have been improved using PISC image compared to conventional method where using PISC image the brightness change in the background and the shading due to moving person are completely removed while only the moving object was extracted. By using our proposed template matching, we successfully track the interest object among the objects appear on the scene. Our identification method using color and spatial feature of the object could extract the moving objects on the successive frame with identification rates more than 90%.

However, our proposed method still has limitations such as to determine the interest object when the other objects come into the scene in the same time. Our algorithm could not determine which object to be the interest object. And also, when the object too small to be tracked, our algorithm could not identify as same object as the interest object. The object similarity is low to be identified. Therefore the identification result is not 100%. The improvement of the methods is necessary to improve the algorithm and speed up the processing time. These are remaining for our future works.

## REFERENCES


[1] Y. LIU, A. Haizho and X. Guangyou, "Moving object detection and tracking based on background subtraction", *Proc. Society of Photo-Optical Instrumentation Engineers (SPIE),* vol. 4554, pp. 62–66, 2001

[2] A. J. Lipton, H. Fujiyoshi and R. S. Patil, "Moving target classification and tracking from real-time video," *Proc. IEEE Workshop Applications of Computer Vision*, pp. 8-14, 1998.

[3] D. Meyer, J. Denzler and H. Niemann, “Model based extraction of articulated objects in image sequences for gait analysis,” *Proc. IEEE Int. Conf. Image Processing*, pp. 78-81, 1998.

[4] C. Wren, A. Azerbaijani, T. Darrell and A. Pentland, ”Pfinder: Real-time tracking of the human body”, *IEEE Trans. on Pattern Anal. and Machine Intell.*, Vol. 19 No.7, pp. 780-785, 1997.

[5] J. W. Davis and A. F. Bobick. “The representation and recognition of action using temporal templates” *Proc. of Computer Vision and Pattern Recognition*, pp. 928-934, 1997.

[6] Y. Satoh, S. Kaneko and S. Igarashi, “Robust Object Detection and Segmentation by Peripheral Increment Sign Correlation Image”, *System and Computer in Japan*, Vol. J 84-D-II, No.12, pp. 2585-2594, 2001.

[7] F. Cheng and Y. Chen, “Real time multiple objects tracking and identification based on discrete wavelet transform”, *Journal of the pattern Recognition Society*, Vol. 39, pp 1126-1139, 2006.

[8] E. Stringa, “Morphological change detection algorithms for surveillance applications”, *Proc. British Machine Vision Conf.*, pp. 402-412, 2000.

[9] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA., 2001.


TABLE I. IDENTIFICATION RESULTS

| **Experiment** | **Identification rates** [%] |
|---|---|
| 1 | 92.8 |
| 2 | 91.9 |
| 3 | 91.6 |

Fig. 5. Tracking the target from two moving persons move in the different direction

Fig. 6. Tracking the target from two moving persons move in the same direction

Fig. 7. Tracking the target object from three moving person emerging in the scene